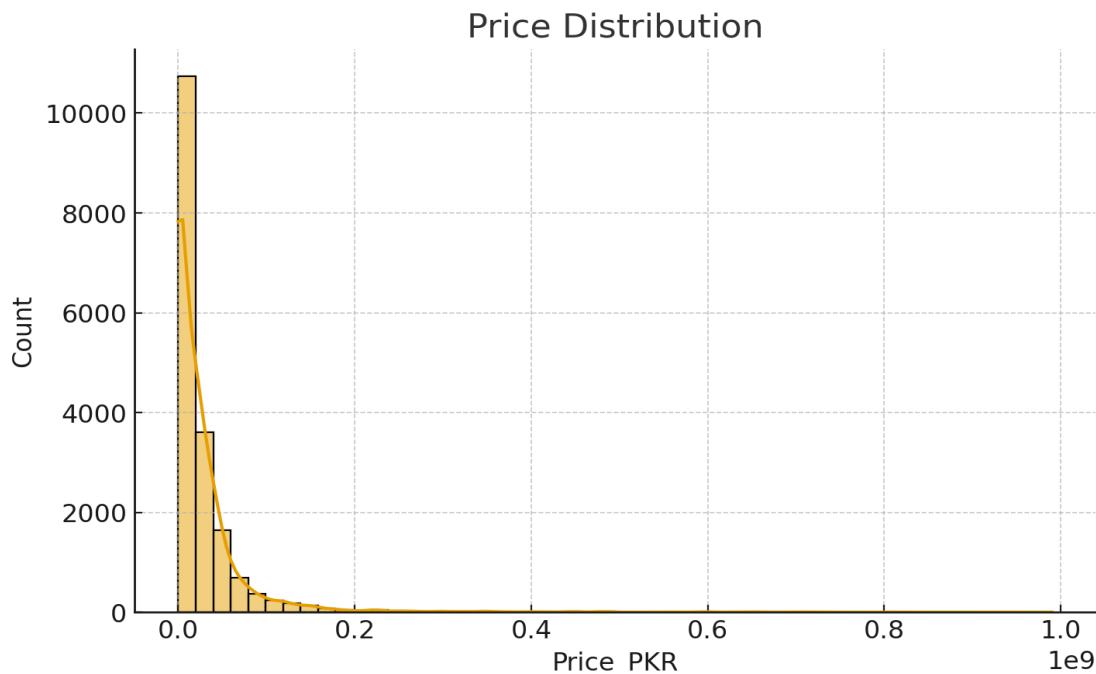
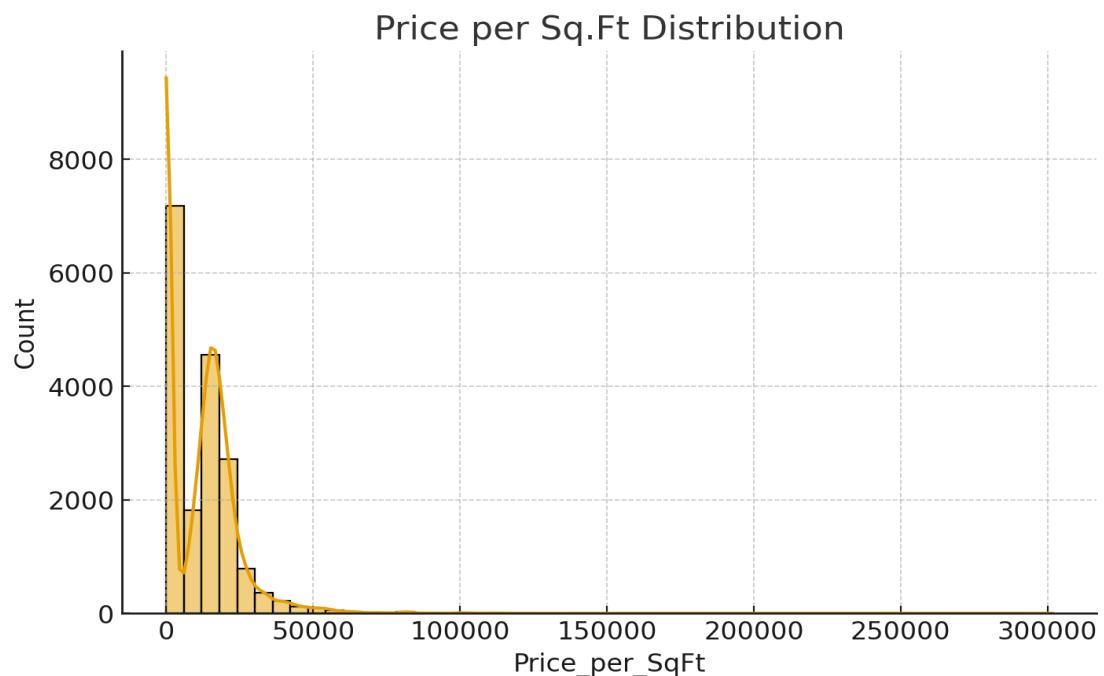


Real Estate Market Insights: Zameen.com Listings in Pakistan Prepared by: Ayesha Jahangir

1. Problem Statement This analysis explores key drivers of residential property prices in Pakistan using scraped listings data from Zameen.com. The aim is to understand how location, property type, covered area, and amenities influence market pricing.
2. Dataset Overview The dataset contains scraped listings including city, area, price, property type, bedrooms, bathrooms, and multiple amenity fields. Several inconsistencies and mixed units required cleaning.
3. Data Cleaning & Preprocessing Removed duplicate entries. Converted PKR prices (Crore/Lakh/Thousand) into numeric PKR. Standardized area units (Marla, Kanal, Sq.Yd → Sq.Ft). Cleaned city labels and whitespace. Converted bedrooms and bathrooms to numeric. Computed Price_per_SqFt.

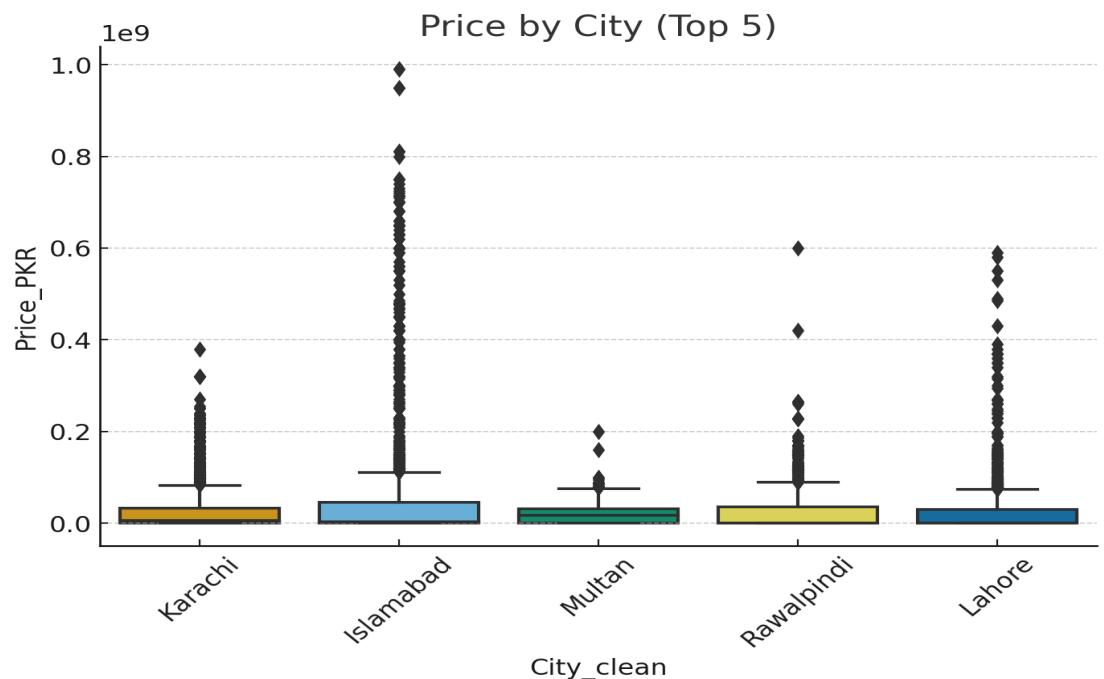


4. Missing Values Core variables were mostly complete. Amenity fields contained many missing values and were used only for descriptive insights.



5. Feature Engineering **Price_per_SqFt**: Key comparison metric. **City_clean**: Standardized city names.

6. Exploratory Findings **City-Level Insights**: Islamabad: Highest median prices. Karachi: Strong demand and PPSF. Lahore & Rawalpindi: Mid-to-upper tier markets. Faisalabad & Multan: More affordable. **General Patterns**: Price distribution is right-skewed. Most houses have 3–6 bedrooms. PPSF is more reliable than total price.



7. Recommendations Invest in Karachi/Islamabad for premium markets. Lahore/Rawalpindi provide balanced entry points. Faisalabad/Multan suitable for budget buyers. Platforms should enforce better listing completeness.

8. Limitations & Next Steps Data represents a snapshot only. Unit definitions vary by region. Future work: rental vs sale analysis, predictive models, geospatial insights. Conclusion By cleaning and analyzing the dataset, the project reveals how location and size drive pricing patterns, offering data-backed guidance for investors.